Charles University in Prague, Faculty of Science
Department of Cell Biology

Study programme: Special Chemical and Biological Programmes
Branch of study: Molecular Biology and Biochemistry of Organisms

Michaela Vaškovičová

Využití RNA afinitních značek v RNA biologii

The Use of RNA Affinity Tags in RNA Biology

Bachelor´s thesis

Supervisor: Doc. Mgr. Petr Svoboda, Ph.D.

Prague, 2013

Prohlášení

Prohlašuji, že jsem bakalářskou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 15.5.2013

Michaela Vaškovičová

Acknowledgements

I would like to thank to my supervisor, Petr Svoboda, for his patience and valuable advices. I would also like to thank to my family and friends for their endless support and understanding.

Contents

Abstract

RNA molecules play essential role in many important processes in living organisms. Messenger RNAs transmit genetic information from the DNA to the cellular sites of protein synthesis. RNAs, which do not encode for proteins, have numerous specialized functions in regulation of gene expression, splicing, transport, localization, translation, and stability of RNAs. For example, ribosomal and transfer RNAs are required for the translation of proteins; small nuclear RNAs are involved in splicing, and small nucleolar RNAs modify other RNAs. MicroRNAs and small interfering RNAs are essential regulators of gene expression. Other non-coding RNAs regulate chromatin and transcription.

RNAs are typically present in ribonucleoprotein (RNP) complexes containing various proteins, which are essential for functions of the complexes. Understanding RNA-protein interactions and composition of RNP complexes is critical for delineating their functions. The common strategy in the study of RNP complexes is tagging the RNA component, in order to isolate, purify and analyze the tagged RNAs with their binding partners. Tagged RNP complexes can be subsequently subjected to further studies.

In the first part of my work, I describe four basic strategies of tagging the RNA, which can be used for study of RNAs and RNP complexes. In the second part, I focus on character, functionality and applications of one of these strategies, which is based on RNA aptamers.

Abstrakt

Molekuly RNA hrajú podstatnú úlohu v mnohých dôležitých procesoch v živých organizmoch. Mediátorové RNA nesú genetickú informáciu z DNA do miesta syntézy proteínov. RNA, ktoré nekódujú proteíny, majú mnoho špecializovaných funkcií v regulácii génovej expresie, zostrihu, transporte, lokalizácii, translácii a stabilite RNA. Ribozomálne a transferové RNA sú nevyhnutné pre transláciu proteínov, malé jadrové RNA sa podieľajú na zostrihu a malé jadierkové RNA modifikujú iné molekuly RNA. MikroRNA a malé interferujúce RNA sú nevyhnutnými regulátormi génovej expresie. Ďalšie nekódujúce RNA regulujú chromatín a transkripciu.

RNA sa obvykle vyskytuje v ribonukleoproteínových (RNP) komplexoch, obsahujúcich rôzne proteíny, ktorých prítomnosť je nevyhnutná pre funkciu daného komplexu. Pochopenie interakcií medzi RNA a proteínmi a tiež zloženie RNP komplexu je kritickým bodom pri charakterizácii ich funkcií. Bežne používanou stratégiou štúdia RNP komplexov je značenie RNA tak, aby bolo možné izolovať, čistiť a analyzovať značené RNA spoločne s ich väzbovými partnermi. Značené RNP komplexy môžu byť následne podrobené ďalšiemu výskumu.

V prvej časti svojej práce popisujem štyri základné stratégie značenia RNA, ktoré môžu byť použité na štúdium RNA a RNP komplexov. V druhej časti bližšie popisujem vlastnosti, funkcie a aplikácie jednej zo stratégií, ktorá je založená na RNA aptaméroch.

1. Introduction

The importance of ribonucleic acid (RNA) became apparent many decades ago, when it was discovered that transmission of genetic information from DNA to proteins is carried out by RNA molecules, including messenger RNA (mRNA), transfer RNA (tRNA), and ribosomal RNA (rRNA). In subsequent research it was found that RNA molecules have many other functions. During the last two decades a large number of non-coding RNAs with various functions was discovered and functionally characterized. Many non-coding RNAs participate in distinct steps of gene expression. Small nuclear RNAs are involved in pre-mRNA splicing, small nucleolar RNAs control chemical modification of RNAs, microRNAs, small interfering RNAs, and long intergenic non-coding RNAs are important regulators of gene expression. Many regulatory RNAs are still to be identified and characterized (reviewed in Eddy, 2001).

RNAs are typically associated with RNA-binding proteins carrying out various functions. Thus, studying functions of RNAs and RNA-binding proteins is a crucial point in the understanding mechanisms important for the cell and the entire organism.

Discovery of new non-coding RNAs is very rapid, and efforts to clarify their functions and activities bring many problems. The key problem in functional analysis of ribonucleoprotein complexes is the difficulty to identify RNA-binding partners, which are involved in their biological functionality (Lee et al., 2013). Thus, numerous RNA affinity purification approaches were developed, which capitalized on the experience from studying protein complexes using protein affinity tags.

Protein affinity tags have irreplaceable role in detection, isolation and purification of proteins. They allow for studying functions and chemical properties of individual proteins or protein complexes. Protein affinity tags bind their targets with high affinity and can be fused with a detection system or can be immobilized on a chromatography resin. An isolated complex can be removed by competitive elution or cleaved off by a protease (Srisawat and Engelke, 2002). Protein affinity tags differ in their properties, and are used in different applications and different biological systems. Among the most common protein affinity tags belong: polyhistidine (His-tag), glutathione S-transferase (GST), protein A, hemagglutinin epitope

(HA), myc epitope, FLAG epitope, Strep-tag and the tandem affinity protein (TAP) tag (reviewed in Walker et al., 2008).

Four basic strategies to tag RNA were developed for the study of RNAs and ribonucleoprotein complexes: (1) chemical modification, (2) hybridization of affinity tagged-oligonucleotides, (3) insertion of a well-characterized protein-binding RNA sequence, and (4) inclusion of an artificially selected RNA motif (aptamer) (reviewed in Walker et al., 2008). These methods have different advantages and disadvantages, which determine their usage. RNA affinity tags can be used for studies of RNA binding motifs or for identification, purification, and characterization of RNA-binding proteins and protein complexes. Another application of RNA affinity tags is functional analysis of RNA-binding proteins.

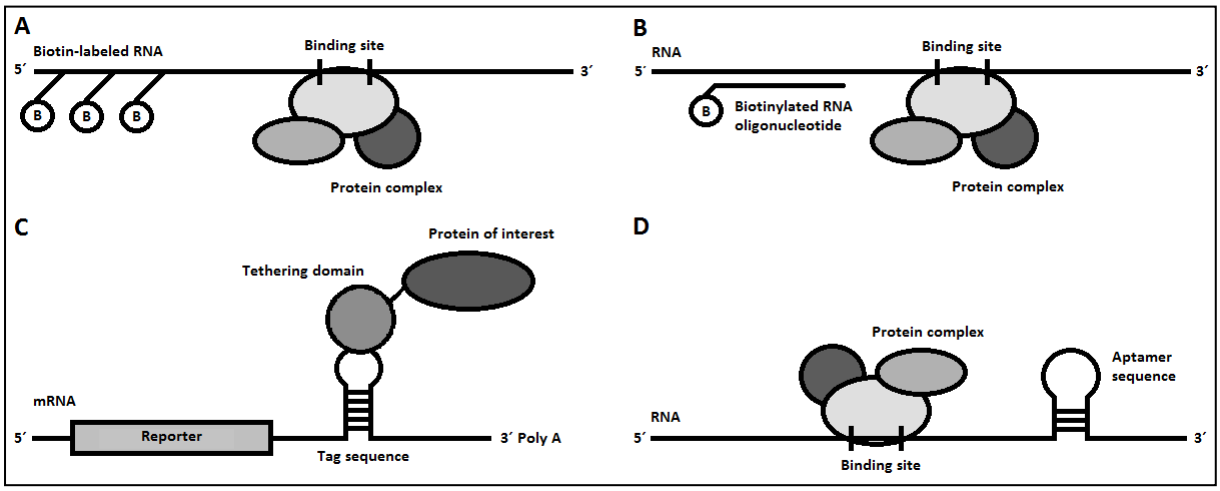

Figure 1 – RNA affinity tags. (A) Chemical tagging, (B) hybridization of affinity tagged-oligonucleotides, (C) incorporation of a well-characterized protein-binding RNA sequence, and (D) incorporation of aptamer.

In the first part of my work, I will describe the principles of the four different tagging strategies and compare their usage. In the second part, I will focus on character, functionality and applications of RNA aptamers.

2. RNA affinity tags

2.1. Chemical tags

RNA can be chemically tagged by incorporating modified ribonucleotide triphosphates containing biotin, digoxigenin, fluorescent dyes, or other compounds. Biotin-labeled RNA can be purified via binding to streptavidin beads, which allows for isolating from cellular extract of ribonucleoprotein complexes or recombinant proteins, which interact with modified RNA (Figure 2). However, there are disadvantages of chemical modifications such as potential alterations of RNA secondary structure, which can interfere with formation of a ribonucleoprotein complex. Biotin-labelling is typically used for *in vitro* studies. Biotin-labelling is not used for *in vivo* studies, because generated complexes may not reflect the native character of ribonucleoprotein complexes (Walker et al., 2008).

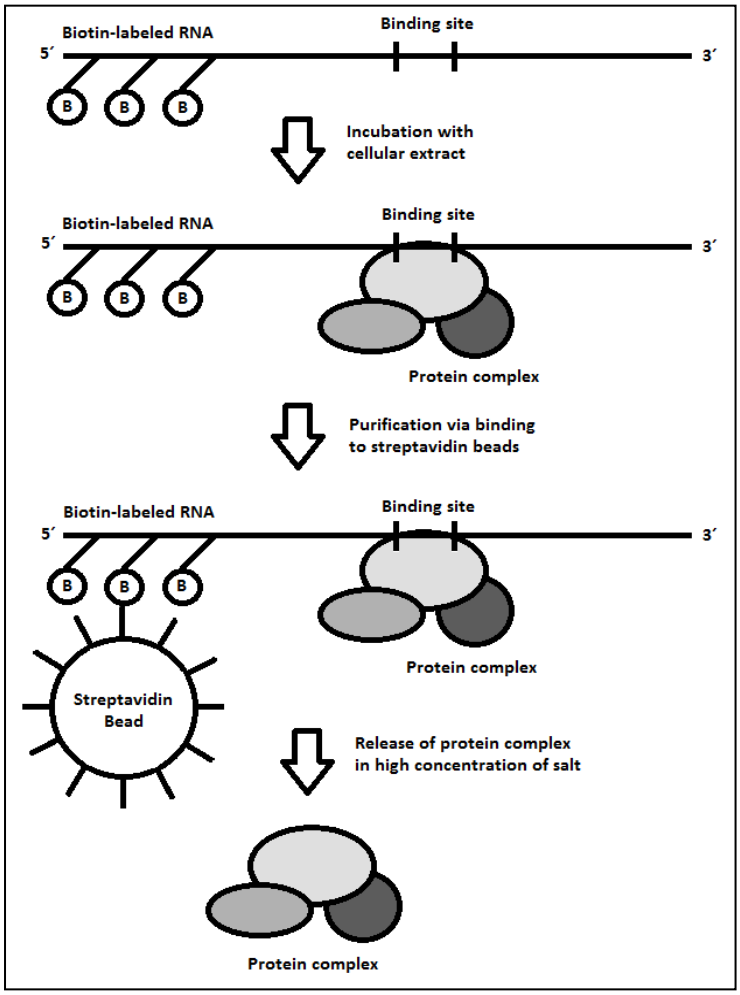

Figure 2 – Isolation of protein complex using a chemical tag. Biotin-labeled RNA is incubated with cellular extract and specific protein complex interacts with binding site; purification of the ribonucleoprotein complex via binding to streptavidin beads is followed by release of protein complex in high concentration of salt.

A typical example of chemical tagging is analysis of iron-responsive element - binding protein (IRE-BP) in human liver (Rouault et al., 1989). Briefly, RNAs carrying iron-responsive sequence (IRE) were transcribed *in vitro*, biotinylated and subsequently incubated with a cellular extract. Binding of ribonucleoprotein complexes to avidin was followed by incubation with biotin-agarose beads and IRE-bound proteins are released by high concentration of salt (Rouault et al., 1989).

2.2. Oligonucleotide tags

Hybridization of affinity tagged oligonucleotides is often used for isolating ribonucleoprotein complexes from cellular extracts. A common way to isolate and purify ribonucleoprotein complexes is to use biotinylated oligonucleotides complementary to single-stranded sequences of RNA components of ribonucleoprotein complexes (Figure 3). Complex release under native conditions can be achieved using a competitor oligonucleotide; other strategies such as high temperature can be used under denaturing conditions (Walker et al., 2008). Using an oligonucleotide tag has many advantages over methods based on incorporation of RNA sequence into the target molecule: (1) accessibility of different sites in the RNA for synthetic oligonucleotide probes can be rapidly tested, (2) foreign sequence should not cause changes in the secondary structure of the RNA *in vivo*, (3) this method can be used when genetic manipulation to tag an RNA is not possible, since the oligonucleotide tag does not include the insertion of the tag sequence into the target RNA sequence (reviewed in Walker et al., 2008). However, the increasing complexity of ribonucleoprotein complex can cause decreasing accessibility of specific sites for oligonucleotides probes, which makes this method inappropriate for study of complex RNPs.

Many different ribonucleoproteins have been isolated by this method. For example, the telomerase from *Euplotes aediculatus* was purified from cellular extract using biotinylated 2´-O-methyl oligonucleotides complementary to the template region of telomerase RNA (Lingner and Cech, 1996). Complex of telomerase and specific oligonucleotides were immobilized on an avidin column. Washing was followed by elution of telomerase under native conditions using displacement oligonucleotides (Lingner and Cech, 1996). Similarly, U2 and U4/U6 small nuclear ribonucleoproteins were purified from HeLa nuclear extract. For study of U4/U6 small nuclear ribonucleoproteins were used biotinylated oligonucleotides

complementary to 3´ terminal domain of U6 RNA or stem 1 region of U6 RNA (Wolff and Bindereif, 1992). Biotinylated oligonucleotides complementary to two separate regions of U2 small nuclear ribonucleoprotein were used for study of U2 small nuclear ribonucleoproteins (Blencowe et al., 1989).

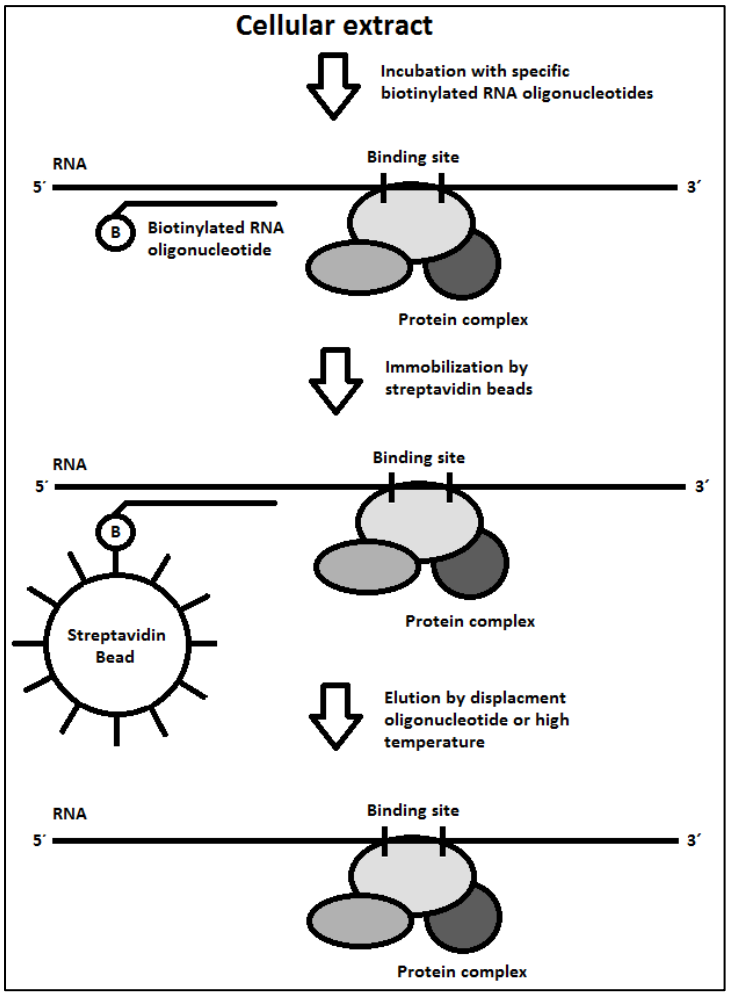

Figure 3 – Isolation of ribonucleoprotein complex using oligonucleotide tag. Cellular extract is incubated with specific biotinylated RNA oligonucleotides; complex is immobilized on avidin column or streptavidin beads and the RNA-protein complex is eluted by displacement oligonucleotides or high temperature.

2.3. Tags based on RNA-protein interaction

Defined RNA sequences from well-characterized RNA-protein interactions can be employed for studying RNA-binding proteins. Tethering assay was developed to simplify studies of RNA-binding proteins, because it allows to separate the function of an RNA-binding protein from mRNA binding (Coller and Wickens, 2002). Tethering assay is based on tethering a protein of interest to a reporter mRNA with high affinity and specificity to a sequence of the reporter through a heterologous RNA-binding domain (Figure 4). First, the fusion protein is created, in which a tethering domain is fused with a protein of interest. Second, an mRNA

reporter (for example LacZ, luciferase) is used, which carries one or more tether binding sites (tag sequences). After coexpression in a cell, the protein of interest binds reporter mRNA through the fused RNA-binding domain and the effects of fusion protein on mRNA stability, localization, and translation of the reporter are analyzed (Coller and Wickens, 2007).

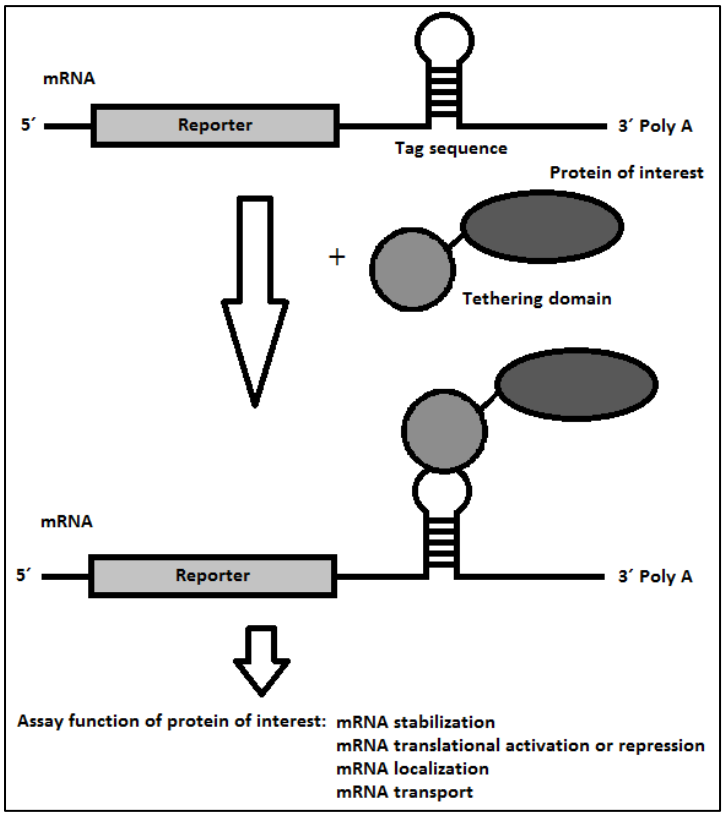

Figure 4 –Tethering assay using the 3´UTR. Tethering domain and protein of interest bind to the mRNA through tag sequence, which is positioned on the 3´UTR of mRNA. The function of protein of interest is subsequently assayed (modified from Coller and Wickens, 2002).

Two peptides of bacteriophage origin are used as common tethers: the MS2 coat protein and lambda N-protein. The MS2 coat protein seems to be used more frequently than lambda N-protein, apparently because of the historical grounds rather than functional differences. Both, the MS2 coat protein and lambda N-protein were shown to be a comparable tool for tethering (Coller et al., 1998; De Gregorio et al., 2001). In addition, spliceosomal U1A protein and iron response element binding protein (IRP) have been used as successful tethers (De Gregorio et al., 1999; Finoux and Seraphin, 2006). The usage of MS2 protein, N-protein, U1A protein or IRP is circumstantial. In principle, all of them can be used for measuring the effect of tethered protein on mRNA metabolism or for purification of protein complexes.

2.3.1. MS2 bacteriophage coat protein

As mentioned above, the MS2 coat protein is the most common tether. MS2 bacteriophage belongs to the group I of bacteriophages, and it is closely related with R17 bacteriophage, which is also used as tether (Keryer-Bibens et al., 2008). MS2 and R17 coat proteins bind to specific sequence elements on mRNA for bacteriophage replicase and repress translation. MS2 and R17 binding sites and binding proteins are highly similar (Keryer-Bibens et al., 2008).

The MS2 and R17 coat proteins are relatively small (14kDa) and biochemistry of their binding to the target sequence is well characterized. The MS2 coat protein binds with high specificity and affinity to a stem-loop motif, which consist of 21 nucleotides (Figure 5A). Only the loop sequence is important for binding (Peabody, 1993). The affinity can be increased by mutations in RNA binding motif; for example, substitution of a single U to a C (at position -5) increases affinity 50-fold over the wild type (Figure 5B) (Lowary and Uhlenbeck, 1987). The MS2 coat protein interacts with its binding site as a dimer, so for every binding site in a reporter mRNA, there are two MS2 proteins bound. Remarkably, the MS2 coat protein has potential to aggregate at concentrations of the MS2 coat protein readily achieved *in vivo* (Johansson et al., 1997). Thus, the specific mutants lacking this potential should be used in tethering assay.

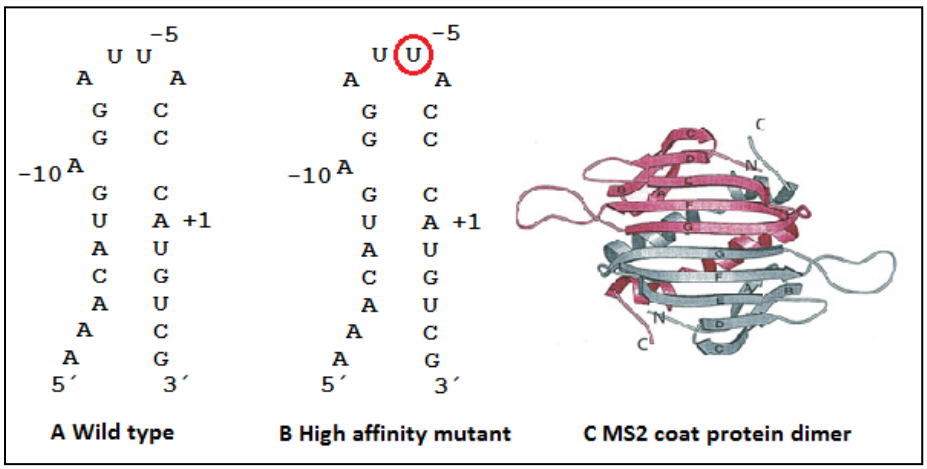

Figure 5 - MS2 RNA stem loop and MS2 coat protein dimer. (A) MS2 stem loop sequence in wild type, (B) MS2 stem loop sequence in high affinity mutant and (C) MS2 coat protein dimer (adapted from Keryer-Bibens et al., 2008).

The first protocol for purification of RNAs and RNA-protein complexes by using the R17 tethering was described by Bardwell and Wickens (Bardwell and Wickens, 1990). The MS2 coat protein was used in a three hybrid system to detect RNA-protein interactions *in vivo* (SenGupta et al., 1996). The use of MS2 tethering assay can be exemplified by a study of the

role of poly(A) binding protein (PAB1P) (Coller et al., 1998). In this case the MS2 binding site was placed in a reporter mRNA 3´UTR and PAB1P fused to MS2 peptide was used for tethering (Coller et al., 1998). The mechanism of stimulation of translation by stem-loop binding protein (SLBP) in vertebrate cells was also examined using tethering assay (Gorgoni et al., 2005).

2.3.2. N-protein

Bacteriophage λ N-protein binds to a specific sequence element in transcribed RNA and has a regulatory function in the process of transcriptional termination (Keryer-Bibens et al., 2008). The N-protein binds an RNA stem-loop structure in the early phage operon called B-box, which consists of 19 nucleotides (Scharpf et al., 2000). Both, the stem and the loop, are important for binding the N-protein, which functions as a monomer, so one protein binds to one binding site (Coller and Wickens, 2007). Only 22 amino acids are required for the interaction of N-protein with B-box with affinity of binding of the full-length protein. The small size of the N-protein is advantageous, since it minimizes potential interference with function of a tethered protein due to steric hindrance (Baron-Benhamou et al., 2004).

The N-protein as a tether was, for example, used to test whether the binding of m^7GpppN cap binding subunit elF4E to mRNA is sufficient to drive productive translation (De Gregorio et al., 2001). Mutant forms of elF4E were fused with RNA-binding domain of the N-protein; fused proteins interacted with a B-box placed in a spacer of a bicistronic reporter mRNA (De Gregorio et al., 2001). The N-protein can be also used for purification of RNAs. To facilitate the release of a ribonucleoprotein complex bound to the N-protein, the ARiBo tag was developed (Di Tomasso et al., 2011). The ARiBo tag contains the *glmS* ribozyme and a B-box (Figure 6). A ribonucleoprotein is purified by immobilization of Glutathione-Sepharose beads via binding to N-protein fused with Glutahione-S-Transferase (GST). The ribonucleoprotein is eluted by self cleavage of *glmS* ribozyme activated by glucosamine-6-phosphate (Di Tomasso et al., 2011).

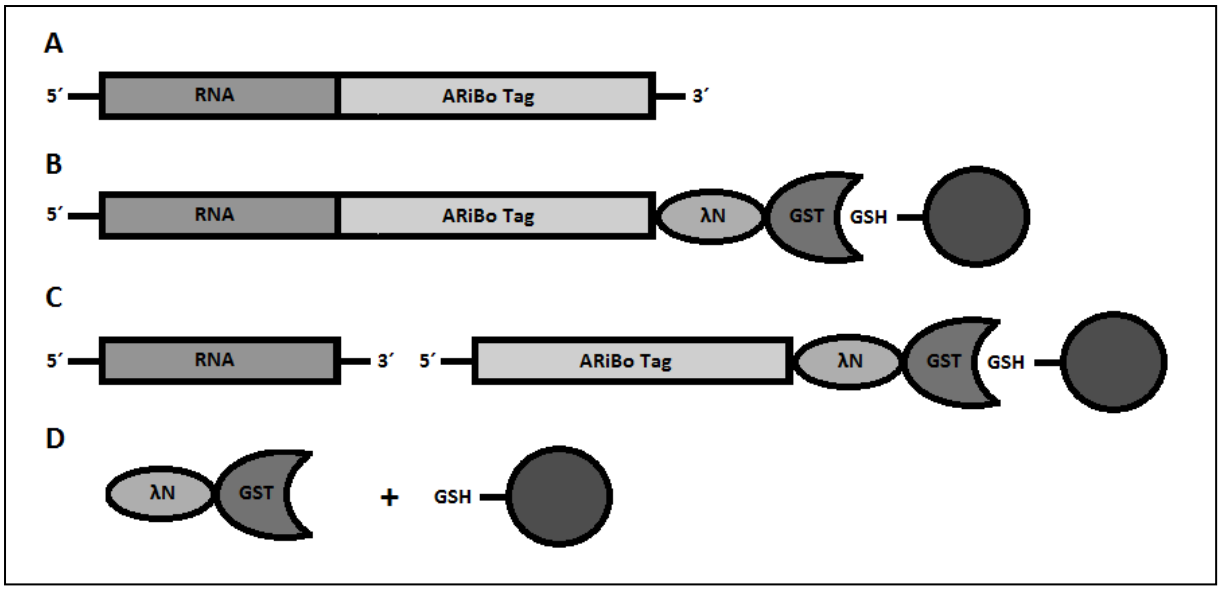

Figure 6 – Affinity purification using the ARiBo tag. (A) *in vitro* transcription of the ARiBo tag fused with RNA of interest, (B) affinity purification via binding to a λN-protein fused with GST and immobilization on Glutathione-Sepharose beads, (C) self cleavage of *gmlS* with glucosamine-6-phosphate and elution of the RNA of interest, and (D) regeneration of affinity resin (modified from Di Tomasso et al., 2011).

2.3.3. U1A protein and IRP

U1A protein and IRP are also functional tethers, but their use is not as widespread as that of MS2 coat protein or N-protein (reviewed in Coller and Wickens, 2007). This is likely due to their larger sizes; U1A protein has 38kDa and IRP has 97kDa. U1A protein is a U1 small nuclear ribonucleoprotein polypeptide A, which binds with a high affinity and specificity to a hairpin consisting of 30 nucleotides (van Gelder et al., 1993). The iron response element binding protein (IRP) binds with high affinity to a 30-nt hairpin (Barton et al., 1990). Both U1A and IRP bind to their binding sites as monomers (Coller and Wickens, 2007).

Tethering based on U1A was used, for example, in functional analysis of the yeast POP2 deadenylase subunit *in vivo* (Finoux and Seraphin, 2006). In this study, the POP2 factor was fused with U1A RNA binding domain and U1A binding sites were inserted into the reporter RNA (Finoux and Seraphin, 2006). Tethering system employing IRP was used, for example, to determine the function of elF4G subunit in ribosome recruitment *in vivo* (De Gregorio et al., 1999). The elF4G fused with IRP-1 was bound to mRNA, which consisted of IRP binding site in a spacer of a bicistronic reporter mRNA (De Gregorio et al., 1999).

2.3.4. Potential problems during functional analysis by protein tethering

There are several potential problems, when tethering assay is used to study effects of a protein tethered to a reporter mRNA. In the following section I will describe some of these potential problems.

Several common controls are important when performing a tethering assay. When using a tethering assay, it is necessary to address: (1) binding of the tethering domain alone, has no effect on the studied RNA, (2) insertion of tethering site has no effect on the studies RNA, and (3) the observed effect is caused by the protein of interest bound to the RNA via the tethering domain (Coller and Wickens, 2007).

When the tethering system is used for functional analysis, only clear effect of the tethered protein on reporter mRNA has informative value and can be interpreted (Keryer-Bibens et al., 2008). Lack of effect can be caused by numerous factors, since some RNA-binding proteins require their natural cognate binding site, specific position of the binding site on an RNA, or another protein bound in the vicinity for their function (Coller and Wickens, 2007).

The position of tethering site depends on the role of a protein in a given RNA metabolism. Thus, the positioning of tethering site must be carefully considered. In the case of mRNAs, placing the tag sequence to the 3´UTR is a common strategy since many sites that regulate mRNA metabolism are found in the 3´UTR (Coller and Wickens, 2002). Moreover, the 3´UTR offers less limitation, than the 5´UTR or the open reading frame (Stripecke et al., 1994; Grskovic et al., 2003). As the effect of a tethered protein may be dependent on the position of the tethering site, it is often needed to perform several experiments with using different positions of tethering sites and evaluate observed effects.

The number of tethering sites may also influence results of a tethering assay. The tethering assay may work well even if only one tethering site is used, but increasing the number of tethering sites can enhance the signal (Coller and Wickens, 2007). For example Pillai et al., who studied transcription repression mediated by microRNAs, observed that the inhibition of protein synthesis was dependent on the number of the hairpins, which tethered N-hAGO2 to the 3´UTR of the mRNA (Pillai et al., 2004). So the increasing number of binding sites can lead to increased number of tethered proteins and also increased efficiency of the tethering assay.

Many different effects of tethered protein on mRNA can be measured by selecting an appropriate reporter mRNA and the configuration of the assay. For measuring translation activity, LacZ, CUP1, and HIS3 mRNAs are often used in the yeast system while luciferase, CAT, and epitope tags are used in metazoans (reviewed in Coller and Wickens, 2007). To determine the effect of tethered protein to the stability of mRNA, MFA2, PGK1 and YAP1 reporters in yeast and β-globin and luciferase in mammalian system are often used (reviewed in Coller and Wickens, 2007).

Geometry of tethered protein can also change the result of tethering assay. The N- and C-ends of each monomer of MS2 dimer are on opposite sides, but they are relatively close to the complementary end of the other monomer (Figure 5C). However, the geometry of tethered protein does not affect the binding capacity of MS2 protein, but may have an effect of the function of the tethered protein (Keryer-Bibens et al., 2008). Also the orientation of fused proteins may cause the change of tethering assay. For example, the two monomers of MS2 dimer are in anti-parallel orientation, which means that tethered proteins are on opposite sides of complex (Ni et al., 1995). The change in flexibility for the two associating monomers could appear, if the monomers are in the parallel orientation (Keryer-Bibens et al., 2008).

2.3.5. Potential problems during purification by protein tethering

Inability to release or elute the RNA-protein complex is one of the problems, which may appear when a tethering system is used for purification of protein complexes. This problem is caused by high binding affinity of tethering domain to the RNA. However, there are some ways to achieve elution of protein or RNA molecule of interest. Fusion of the maltose binding protein (MBP) with tethering domain allows elution of a protein of interest using maltose as binding competitor (Das et al., 2000). Another possibility is to insert a protease cleavage site between the tethering domain and studied protein, which allows for releasing the studied protein upon specific protease cleavage (Leonov et al., 2003). Batey et al. developed a rapid method for purification of RNAs from *in vitro* transcription reactions (Batey and Kieft, 2007). This method is based on incorporation of dual function tag to target RNA during *in vitro* transcription. The tag consists of two elements: (1) *glmS* ribozyme, from *Bacillus subtilis*, activated by glucosamine-6-phosphate and (2) hexahistidine-tagged form of the MBP-MS2 coat fusion protein (Batey and Kieft, 2007). This protein can be immobilized using Ni^{2+}-affinity column. After immobilization, the tagged RNA is eluted by adding glucosamine-6-

phosphate, which activates the ribozyme (Batey and Kieft, 2007). This new method allows rapid parallel purification of RNAs.

Taken together, the tethering assay allows studying, comparing and determining different roles of RNA-binding proteins in mRNA metabolism. Applications of this assay are widespread. Tethering assay can help us in defining functions of tethered proteins and also in clarifying their effects on the metabolism of mRNA (Coller and Wickens, 2007).

2.4. Aptamer tags

Aptamers are single-stranded nucleic acid molecules, which show a high affinity and specificity to small molecules or complex macromolecules. Aptamers are characterized by specific three-dimensional structure and they are developed *in vitro* in a process called SELEX (systematic evolution of ligands by exponential enrichment) (Ellington and Szostak, 1990; Tuerk and Gold, 1990). In general, the aptamers are used in many different areas of research. Because of the versatility of aptamers, they can be used as tools in biotechnology, diagnostics and therapeutics; but also as biosensors, or anti-infectious agents (reviewed in Bunka and Stockley, 2006). In RNA biology, they are used as tags for affinity purification of RNAs and ribonucleoprotein complexes (reviewed in Walker et al., 2008). In the following section I will describe the using of aptamers for studies of RNAs and ribonucleoprotein complexes.

3. Aptamers for studies of RNAs and RNA complexes

Aptamers are single-stranded RNA molecules used for purification of RNAs and ribonucleoprotein complexes. This method relies on insertion of an aptamer sequence into a gene. The isolation of aptamer-tagged RNAs and RNA-protein complexes from cellular extract subsequently consists of several steps: (1) incubating the cellular extract with affinity matrix (binding tagged RNAs to affinity matrix), (2) washing step (removal of unbound RNAs and other compounds of cellular extract), (3) elution step (elution of tagged RNAs from affinity matrix) and (4) further characterization of tagged RNAs and RNA-protein complexes (Figure 6).

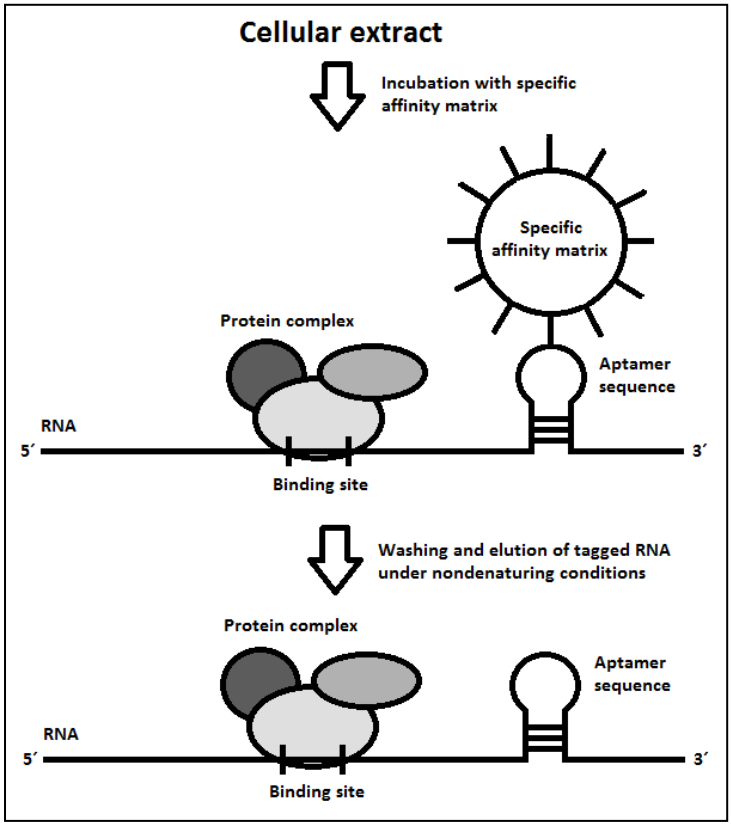

Figure 6 – Isolation and purification of RNAs and ribonucleoprotein complexes using aptamers. Cellular extract is incubated with specific affinity matrix; only the aptamer-tagged RNAs interact with affinity matrix. Removal of unbound RNAs is followed by elution of tagged RNAs under nondenaturing conditions.

RNA aptamers and their target molecules, which are designed for purification and isolation of RNAs and ribonucleoprotein complexes, should satisfy several criteria: (1) high affinity and specificity for a potential affinity resin, (2) possibility of elution of the complex under native conditions, (3) minimal affinity of the potential affinity resin to nonspecific RNAs, (4) acceptable price, and (5) availability of the potential affinity resin. Streptavidin and dextran B512 (in the insoluble form of Sephadex beads) satisfy these criteria, so the

streptavidin aptamer and Sephadex aptamer were developed (Srisawat and Engelke, 2001; Srisawat et al., 2001). These two aptamers have different strengths and weaknesses (summarized in Table 1), which determine their use under different conditions and circumstances.

Table 1 – Comparison between two RNA affinity tags (modified from Srisawat and Engelke, 2002)

Aptamer	**D8 Sephadex RNA motif**	**S1 Streptavidin RNA motif**
Strengths	-cheap affinity resin -concentration of ligand on the beads is nearly infinite -practical purification from large starting quantities -elution with denaturants or by competition with dextran B512	-availability of affinity resin from multiple commercial sources -high affinity for streptavidin -no affinity for egg white avidin -elution with biotin is quick and clear -binding is stable under high salt concentrations
Weaknesses	-lower affinity than streptavidin tag -slow loss of bound RNA during washing -difficult removal of dextran from eluate	-expensive affinity resin -lower number of binding sites -need of blocking with egg white avidin

3.1. Sephadex binding aptamer

RNA aptamers, which bind Sephadex with high affinity and specificity, were developed using *in vitro* process called SELEX. Advantages of using Sephadex as a binding matrix are (1) availability, since Sephadex is common matrix used in gel filtration, (2) low price, (3) stability, and (4) possible repeated use. Sephadex mostly consists of repeating units of glucose linked by α-1,6 glucosidic bonds, which gives Sephadex a large number of binding sites for an aptamer (Srisawat et al., 2001). The elution of the aptamer complex from Sephadex can be done under nondenaturing conditions with dextran B512, which can compete efficiently with the aptamer (Srisawat and Engelke, 2002). These features allow for using Sephadex beads and Sephadex aptamers for RNA affinity purification of tagged RNAs from a complex population of cellular RNAs.

3.1.1. Sephadex binding motif

Srisawat et al. selected Sephadex binding aptamers from an RNA library, which had complexity of ~7 x 10^{16} different sequences (Srisawat et al., 2001). The selection lasted 11 rounds and the final sequences were cloned into plasmids. After sequence analysis of selected aptamers that showed affinity to the Sephadex, it was found that these sequences can be

divided into two groups based on their consensus sequences (Figure 7). All aptamers bound Sephadex with high affinity; the highest affinity was found for aptamer D8 (Srisawat and Engelke, 2002). Thus, D8 aptamer was selected for further experiments and more detailed analysis. It was found that a minimal 33 nucleotides-long D8 binding motif (5´ UCCGAGUAAUUUACGUUUUGAUACGGUUGCGGA 3´) is able to bind Sephadex with a similar capacity as the full-length aptamer, consisting of 84 nucleotides (Figure 8) (Walker et al., 2008).

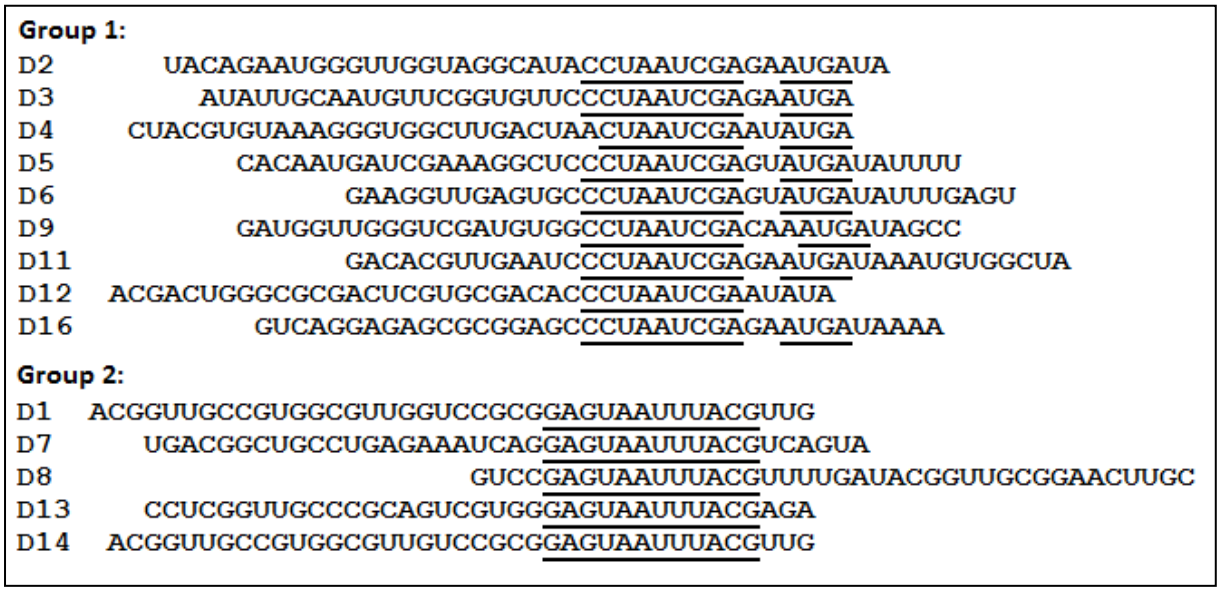

Figure 7 – Classification of Sephadex aptamers (adapted from Srisawat et al., 2001).

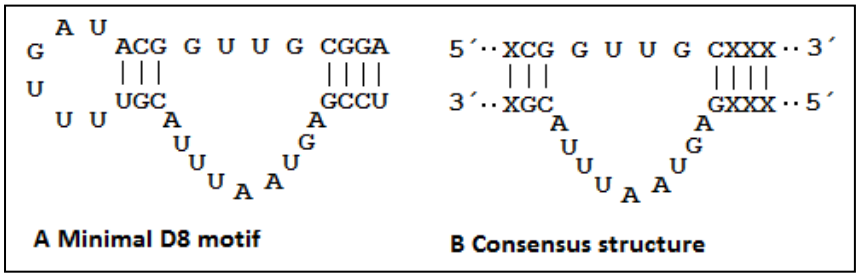

Figure 8 – Structure of D8 aptamer. (A) Minimal binding motif and (B) consensus sequence of D8 aptamer. X represents nonconserved nucleotides (adapted from Walker et al., 2008).

3.1.2. Binding of D8 aptamer to Sephadex

Sephadex beads are formed by cross-linking of dextran B512 with epichlorohydrin (Misaki et al., 1980) and different pore size is achieved by changing the degree of cross-linking. Sephadex G-100 (diameter 40-120 μm) was originally used for the selection of D8 aptamer, but it was found that Sephadex G-200 (diameter 40-120 μm) is a better choice (Srisawat et al., 2001). It is supposed that it is caused by high efficiency of binding larger complexes than by Sephadex G-10 – G-50, which have smaller pore size (Srisawat et al., 2001). At the same

time, the elution of the aptamer complex from Sephadex under nondenaturing conditions with dextran B512 is highly efficient.

Isomaltose, isomaltotriose and isomaltotetraose, oligosaccharides structurally similar to dextran B512 and Sephadex, consist of two, three and four glucose residues that are linked via α-1,6 glucosidic bonds. However, they cannot be used instead of Sephadex for binding D8 aptamer, although their structure is very similar to Sephadex (Srisawat et al., 2001). Since the D8 did not bind isomaltose, isomaltotriose or isomaltotetraose, it was assumed that the optimal binding site of D8 aptamer consist of more than four glucose residues. This characteristic helps to prevent the interference of D8 aptamer with related mono- and oligosaccharides, which are normally present in cellular extract (Srisawat et al., 2001).

The affinity of D8 aptamer to other supporting matrices was also studied, since the ability of D8 aptamer to bind other supporting matrices could cause losses of tagged RNAs and ribonucleoprotein complexes, when additional purification steps are employed. Srisawat et al. compared the affinity of D8 aptamer to the Sephadex with other supporting matrices; Sepharose, Sephacryl, cellulose and pustulan were used (Srisawat et al., 2001). Sepharose, Sephacryl and cellulose are commonly used as supporting matrices in purification applications. However, D8 aptamer showed high affinity and specificity only to Sephadex; binding to another supporting matrices was showed to be ineffective (Srisawat and Engelke, 2002). This is important because acquisition of the ribonucleoprotein complex, consisting of all compounds, sometimes desires to use additional purification steps, which may involve different supporting matrices (Srisawat et al., 2001).

3.2. Streptavidin binding aptamer

Streptavidin aptamers, which bind streptavidin with high specificity and affinity, were also developed by SELEX. Streptavidin is a homotetrameric protein produced by bacteria *Streptomyces avidinii* (Green, 1990). The natural ligand of streptavidin is *d*-biotin and the interaction between streptavidin and *d*-biotin is one of the strongest known noncovalent interactions (Green, 1990). Four moles of biotin bind to one mole of streptavidin with very high affinity, allowing for the elution of aptamer from streptavidin under nondenaturing conditions using biotin as competitor. Similarly to Sephadex, streptavidin is easily available

and is obtainable in a pure form, or conjugated with supporting matrices and fluorescent dyes. It was shown that streptavidin aptamers can be successfully used for detection and purification of RNAs and ribonucleoproteins, while their structures and functions remain unchanged (Srisawat and Engelke, 2001).

3.2.1. Streptavidin binding motif

Streptavidin aptamers were developed by SELEX using an RNA library consisting of more than 10^{16} sequences. The library was subjected to six rounds of selection and amplification to obtain RNA aptamers that bind streptavidin conjugated to agarose beads (Srisawat and Engelke, 2001). Identified aptamers were further selected in three rounds using electrophoretic mobility shift assay (Srisawat and Engelke, 2001). Srisawat et al. have developed eight RNA aptamers binding to streptavidin and have determined their structure (Srisawat and Engelke, 2001). Aptamers were divided into two groups based on the consensus structure (Figure 9). Six of aptamer belong to group 1. S1 aptamer shows the highest affinity and it also produces a specific secondary structure (Figure 10), so S1 aptamer was chosen as a representative streptavidin aptamer for further characterization. S1 streptavidin-binding motif consists of 44 nucleotides
(5′ ACCGACCAGAAUCAUGCAAGUGCGUAAGAUAGUCGCGGGCCGGG 3′). The minimal S1 motif and consensus structure are shown in Figure 10.

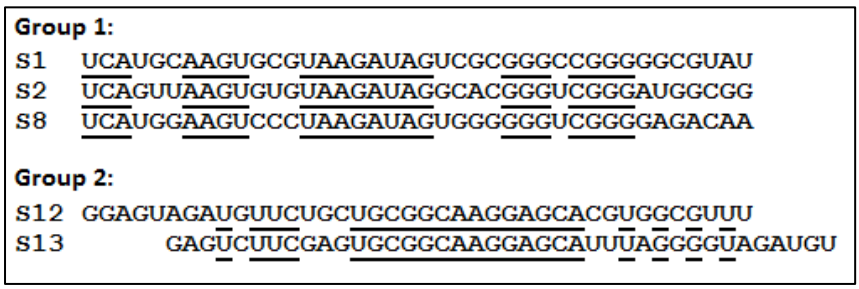

Figure 9 – Classification of streptavidin aptamers (adapted from Srisawat and Engelke, 2001).

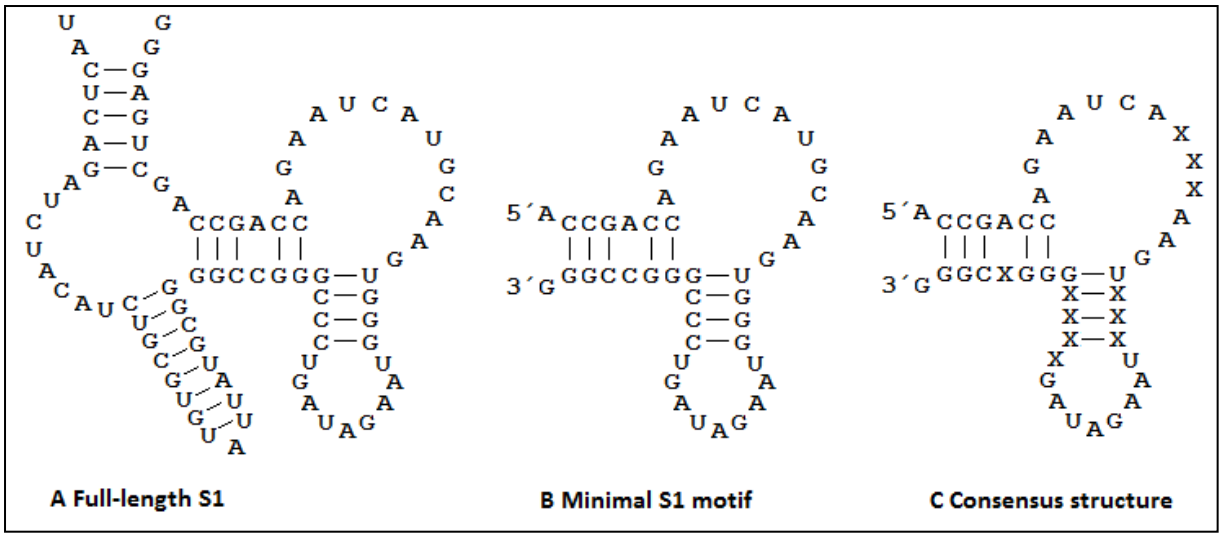

Figure 10 – Structure of S1 aptamer. (A) Full-length S1 aptamer, (B) minimal S1 motif and (C) consensus sequence of S1 aptamer. X represents nonconserved nucleotides (adapted from Walker et al., 2008).

3.2.2. Binding of S1 aptamer to streptavidin

S1 aptamer binds to streptavidin with high affinity, but it can be released from streptavidin under native conditions with biotin. Affinity of interaction between biotin and streptavidin is extraordinarily high, which allows for rapid and complete elution of S1 aptamer from streptavidin in the presence of biotin. The exact mechanism of competition between S1 aptamer and biotin is still unclear, but there are two possible explanations: (1) binding site of aptamer is very close or is identical to the biotin binding pocket, what causes the competition between biotin and aptamer; or (2) aptamer could not recognize and bind the streptavidin, because of the conformation changes generated by the binding of biotin to streptavidin (Srisawat and Engelke, 2001). There are numerous advantages of using S1 aptamer for purification RNAs and ribonucleoprotein complexes. Four key advantages are described below:

- Denaturing or extreme condition, such as low pH, presence of denaturing reagent or high concentration of ions, are necessity for many commonly used techniques for elution of affinity tags. These conditions cause irreversible changes in structure, organization and function of target molecule. However, elution of S1 aptamer from streptavidin can be done under mild conditions by adding biotin to binding buffer and the character of target molecule stay intact, which is one of the greater advantages of using biotin as elution reagent (Srisawat and Engelke, 2001).

- The interaction between biotin and streptavidin seems to be irreversible, which provides protection against rebinding the aptamer back to streptavidin (Srisawat and Engelke, 2002). However, Holmberg et al. showed that biotin-streptavidin interaction can be disrupted by short incubation in non-ionic aqueous solution at temperature above 70°C. In this case, the streptavidin and biotin remain intact and the both molecules can be used multiple times (Holmberg et al., 2005).

- Biotin has small size (244 Da). This feature allows for removing biotin from eluted solution by dialysis or ultrafiltration (Srisawat and Engelke, 2002), if the presence of biotin in the eluate could affect results of further experiments.

- S1 aptamer as affinity tag does not bind egg white avidin. Egg white avidin is protein, which binds to biotin with affinity similar to streptavidin (Green, 1990). This can be very useful when purifying RNAs and ribonucleoproteins from cellular extracts. Cells contain free biotin and bionylated proteins, which bind irreversibly to streptavidin and block the binding of S1 aptamer. This problem can be solved by incubating the extract with avidin, before the binding step during purification, which results in sequestering biotin and avidin (Srisawat and Engelke, 2002).

3.3. Insertion of aptamer tag into the RNA

Insertion of an aptamer tag into the target RNA is one of the essential steps in isolation and purification of RNAs and ribonucleoprotein complexes. There are three main factors to consider when inserting an aptamer tag: (1) folding, (2) steric hindrance, and (3) maintaining the aptamer tag on RNA during purification (Srisawat and Engelke, 2002).

3.3.1. Folding

The aptamer tag must be placed in such way that both, the aptamer tag and RNA, are able to fold correctly (Srisawat and Engelke, 2002). Structural information of RNA is very useful, and predicted secondary structure offers suitable conditions for designing appropriate tagged-RNA constructs (Walker et al., 2008). However, it is often necessary to generate and test several of them. It is also important to test whether the natural biological functions of tagged-RNA are maintained, and whether the tagged-RNA is still able to bind the affinity matrix

(Srisawat and Engelke, 2002). The aptamer tag can be inserted into positions inside an RNA molecule in appropriate position, preferentially in a terminal loop of a long stem, which protrude into solution (Figure 11), or at 5′ or 3′ ends of an RNA molecule (Walker et al., 2008).

3.3.2. Steric hindrance

The steric blockage problem occurs when the tag sequence obstructs the formation of a ribonucleoprotein complex, or when the function of aptamer is obstructed by formation of a ribonucleoprotein complex (Walker et al., 2008). It is difficult to predict whether the steric blockage problem will arise, so the information about the structure of the RNA of the interest is beneficial. Insertion into a protruding stem in a predicted structure may prevent the generation of steric blockage, in the case that precise information about structure is not available (Srisawat and Engelke, 2002).

Another strategy to avoid the steric hindrance is placing a spacer region between the aptamer sequence and the main body of RNA (Walker et al., 2008). By introducing a flexible or rigid linker or by changing the length of spacer (Figure 11), it is possible to influence the interaction of the aptamer with target RNA (Walker et al., 2008). A test of folding is necessary to ensure that no interference occurs between the spacer and the aptamer sequence or the main body of RNA (Srisawat and Engelke, 2002).

3.3.3. Maintaining the aptamer tag on the RNA

Stability of the aptamer tag in an RNA is another problem. In general, protein-based affinity tags are more stable than RNA affinity tag, as RNA aptamers are more sensitive to degradation by nucleases (Walker et al., 2008). The solution of this problem depends on the target RNA, expression system and positioning of aptamer tag. There are two general strategies to protect the aptamer sequence. (1) One of the strategies is placing a strong stem at the 3′ end of the transcript to avoid the degradation of aptamer sequence by 3′-5′ exonucleases, whether the aptamer tag is positioned inside the RNA or at the end of RNA (Srisawat and Engelke, 2002). (2) Another strategy to minimize tagged RNA degradation is possible when the secondary structure of the RNA is available (Srisawat and Engelke, 2002). This strategy is based on placing the aptamer sequence at the terminal loop of an internal stem accessible in solution, which shows no function. In addition, insertion of spacers and flexible

linkers should be considered, because their presence could bring potential risk of degradation by endonucleases (Walker et al., 2008).

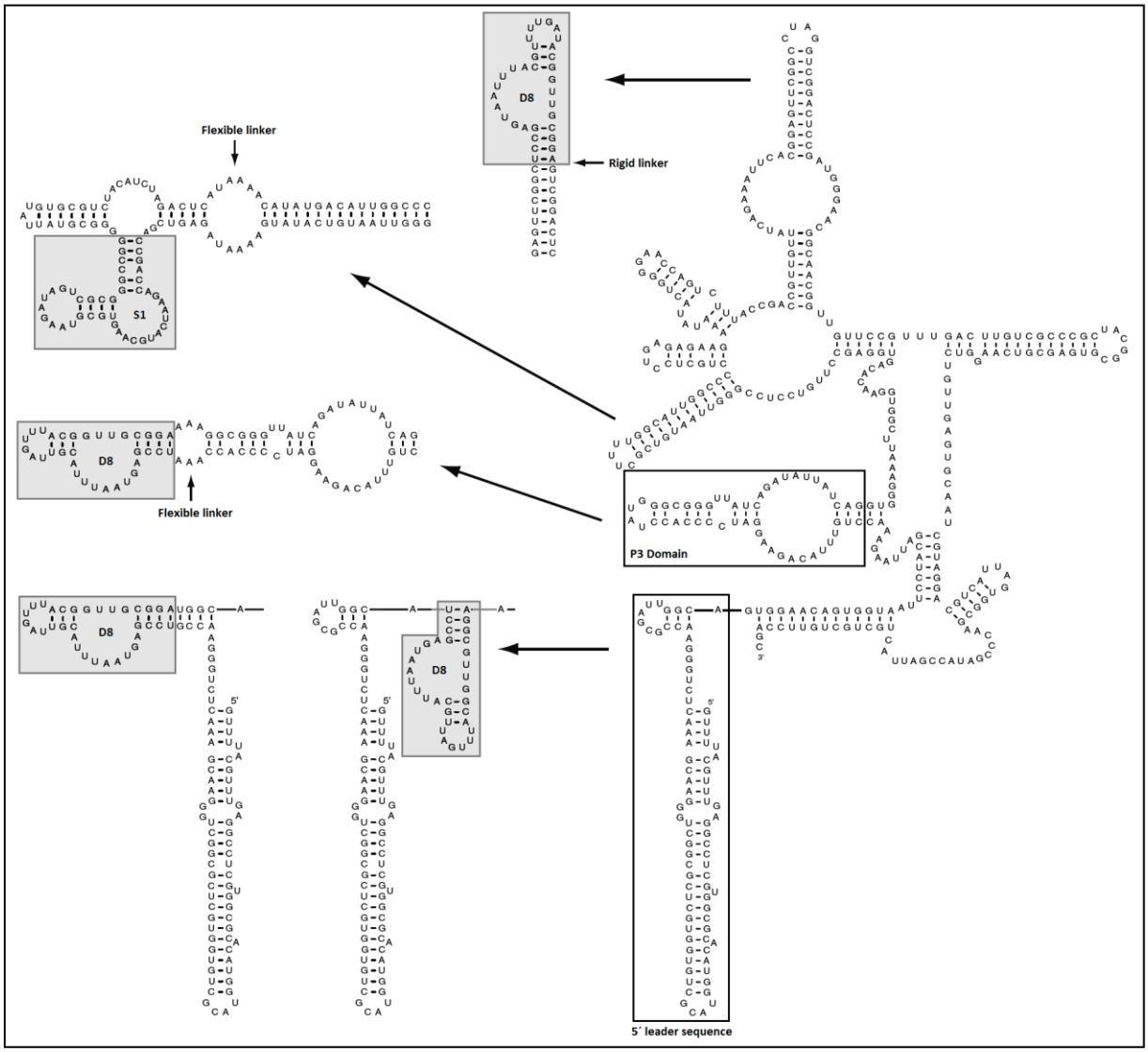

Figure 11 – Different possibilities of insertion of S1 and D8 aptamers into the RNA subunit of ribonuclease P. Minimal aptamer sequences are shown in grey boxes (modified from Walker et al., 2008).

3.4. Application of S1 and D8 aptamers

Both the Sephadex and streptavidin binding aptamers have been successfully used, for isolating either a tagged RNA from total cellular RNA or ribonucleoprotein complexes from cellular extracts (Srisawat and Engelke, 2001; Srisawat et al., 2001). The D8 and S1 aptamer have been used, for example, for study and purification of ribonuclease P (RNase P) from *Saccharomyces cerevisiae* and human S3 cells (Li and Altman, 2002; Srisawat et al., 2002).

RNase P is the endoribonuclease responsible for cleavage of 5′ leader sequences of precursor tRNAs to generate the mature 5′ ends of tRNAs (Frank and Pace, 1998). RNase P is an essential enzyme present in all three domains of life, and it differs in complexity across the phylogenetic domains. The eukaryotic nuclear RNase P shows high complexity, for example, the nuclear RNase P in *Saccharomyces cerevisiae* consists of one RNA subunit and at least nine protein subunits (Houser-Scott et al., 2001; Xiao et al., 2002).

The organization of *S.cerevisiae* nuclear RNase P was studied using aptamer sequences (Srisawat et al., 2002). RPR1 RNA, the RNA subunit of RNase P, had been purified from cellular extract using either the S1 aptamer plus the D8 aptamer, or the S1 aptamer alone (Figure 11) (reviewed in Srisawat and Engelke, 2002). The S1 aptamer was incorporated into region in RPR1 RNA, which was shown to be exposed to the solvent (Figure 11). The S1-tagged RPR1 gene was placed into yeast plasmid and transformed into yeast strain (Srisawat and Engelke, 2001). After expression the crude extract was prepared and incubated with streptavidin-agarose beads. The unbound RNAs were removed by extensive washing. The S1-tagged RPR1 RNA was eluted by adding biotin. The D8 aptamers were incorporated into two regions on RPR1 RNA: (1) 5′leader site, and (2) terminal loop of P3 hairpin domain (Figure 11) (Srisawat et al., 2002). The D8-tagged RPR1 gene was expressed from a yeast plasmid. Prepared crude extracts were incubated with Sephadex G-200 beads. After binding, the Sephadex beads were washed, and the D8-tagged RPR1 RNA was eluted by adding soluble dextran (Srisawat et al., 2002). Aptamer tagged RPR1 RNA was subjected to further analysis of its organization and activity.

It was shown, that S1 aptamer can be also used in mammalian system (Li and Altman, 2002). The S1-tagged human nuclear RNase P can be immobilized from crude extract using streptavidin-agarose beads and eluted by adding biotin. S1 aptamer had also been applied to study the function of protein binding domain (P3) of the RNase P RNA from Hela cells (Li and Altman, 2002).

4. Conclusions & outlook

RNA molecules, together with RNA-binding proteins, take place in regulation of many cellular pathways. Thus, the understanding of the RNAs and ribonucleoprotein complexes functions and mechanism of their regulation is very important for wide range of areas in biology.

RNA affinity tags play indispensable roles in isolation, purification and identification of RNA molecules and ribonucleoprotein complexes. Here, I reviewed four basic strategies to tag RNA molecules. Their choice is dependent on the advantages and disadvantages of each strategy. Chemical and oligonucleotide tagging are commonly used for isolation and purification of RNA-binding proteins and ribonucleoprotein complexes. Tagging based on well characterized RNA-protein interaction is used for studying functions of RNA-binding proteins. Using of aptamers is relatively reliable strategy for isolation of RNAs and ribonucleoprotein complexes under native conditions. However, the selection of specific strategy is dependent on selected RNA or protein of interest, availability of material, and of course, on the aim of the study.

High expenses and technical challenges associated with these techniques stimulate scientists to improve these methods and broaden their applicability. At the same time, new strategies allowing simple and effective purification and identification of ribonucleoprotein complexes are developed. For example, a novel technology using clustered regularly interspaced short palindromic sequences (CRISPR) endoribonuclease was recently described (Lee et al., 2013). This strategy is based on an engineered version of the Csy4 endoribonuclease, which has small size, and high affinity to a short hairpin sequence present in the RNA of interest. Advantage of the mutated Csy4 endoribonuclease is activation of its site-specific RNA cleavage activity by adding imidazole. This strategy allows for rapid purification of tagged RNAs together with their specific binding proteins.

Therefore, while the methods reviewed in my thesis will be still used in the future, it is foreseeable that various novel methods will be added to the existing repertoire.

5. References

Bardwell, V. J. and M. Wickens (1990). "Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method." Nucleic acids research 18(22): 6587-6594.

Baron-Benhamou, J., N. H. Gehring, A. E. Kulozik and M. W. Hentze (2004). "Using the lambdaN peptide to tether proteins to RNAs." Methods in molecular biology 257: 135-154.

Barton, H. A., R. S. Eisenstein, A. Bomford and H. N. Munro (1990). "Determinants of the interaction between the iron-responsive element-binding protein and its binding site in rat L-ferritin mRNA." The Journal of biological chemistry 265(12): 7000-7008.

Batey, R. T. and J. S. Kieft (2007). "Improved native affinity purification of RNA." RNA 13(8): 1384-1389.

Blencowe, B. J., B. S. Sproat, U. Ryder, S. Barabino and A. I. Lamond (1989). "Antisense probing of the human U4/U6 snRNP with biotinylated 2'-OMe RNA oligonucleotides." Cell 59(3): 531-539.

Bunka, D. H. and P. G. Stockley (2006). "Aptamers come of age - at last." Nature reviews. Microbiology 4(8): 588-596.

Coller, J. and M. Wickens (2002). "Tethered function assays using 3' untranslated regions." Methods 26(2): 142-150.

Coller, J. and M. Wickens (2007). "Tethered function assays: an adaptable approach to study RNA regulatory proteins." Methods in enzymology 429: 299-321.

Coller, J. M., N. K. Gray and M. P. Wickens (1998). "mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation." Genes & development 12(20): 3226-3235.

Das, R., Z. Zhou and R. Reed (2000). "Functional association of U2 snRNP with the ATP-independent spliceosomal complex E." Molecular cell 5(5): 779-787.

De Gregorio, E., J. Baron, T. Preiss and M. W. Hentze (2001). "Tethered-function analysis reveals that elF4E can recruit ribosomes independent of its binding to the cap structure." RNA 7(1): 106-113.

De Gregorio, E., T. Preiss and M. W. Hentze (1999). "Translation driven by an eIF4G core domain in vivo." The EMBO journal 18(17): 4865-4874.

Di Tomasso, G., P. Lampron, P. Dagenais, J. G. Omichinski and P. Legault (2011). "The ARiBo tag: a reliable tool for affinity purification of RNAs under native conditions." Nucleic acids research 39(3): e18.

Eddy, S. R. (2001). "Non-coding RNA genes and the modern RNA world." Nature reviews. Genetics 2(12): 919-929.

Ellington, A. D. and J. W. Szostak (1990). "In vitro selection of RNA molecules that bind specific ligands." Nature 346(6287): 818-822.

Finoux, A. L. and B. Seraphin (2006). "In vivo targeting of the yeast Pop2 deadenylase subunit to reporter transcripts induces their rapid degradation and generates new decay intermediates." The Journal of biological chemistry 281(36): 25940-25947.

Frank, D. N. and N. R. Pace (1998). "Ribonuclease P: unity and diversity in a tRNA processing ribozyme." Annu Rev Biochem 67: 153-180.

Gorgoni, B., S. Andrews, A. Schaller, D. Schumperli, N. K. Gray and B. Muller (2005). "The stem-loop binding protein stimulates histone translation at an early step in the initiation pathway." RNA 11(7): 1030-1042.

Green, N. M. (1990). "Avidin and streptavidin." Methods in enzymology 184: 51-67.

Grskovic, M., M. W. Hentze and F. Gebauer (2003). "A co-repressor assembly nucleated by Sex-lethal in the 3'UTR mediates translational control of Drosophila msl-2 mRNA." The EMBO journal 22(20): 5571-5581.

Holmberg, A., A. Blomstergren, O. Nord, M. Lukacs, J. Lundeberg and M. Uhlen (2005). "The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures." Electrophoresis 26(3): 501-510.

Houser-Scott, F., W. A. Ziehler and D. R. Engelke (2001). "Saccharomyces cerevisiae nuclear ribonuclease P: structure and function." Methods Enzymol 342: 101-117.

Johansson, H. E., L. Liljas and O. C. Uhlenbeck (1997). "RNA recognition by the MS2 phage coat protein." Seminars in Virology 8(3): 176-185.

Keryer-Bibens, C., C. Barreau and H. B. Osborne (2008). "Tethering of proteins to RNAs by bacteriophage proteins." Biology of the cell / under the auspices of the European Cell Biology Organization 100(2): 125-138.

Lee, H. Y., R. E. Haurwitz, A. Apffel, K. Zhou, B. Smart, C. D. Wenger, S. Laderman, L. Bruhn and J. A. Doudna (2013). "RNA-protein analysis using a conditional CRISPR nuclease." Proceedings of the National Academy of Sciences of the United States of America 110(14): 5416-5421.

Leonov, A. A., P. V. Sergiev, A. A. Bogdanov, R. Brimacombe and O. A. Dontsova (2003). "Affinity purification of ribosomes with a lethal G2655C mutation in 23 S rRNA that affects the translocation." The Journal of biological chemistry 278(28): 25664-25670.

Li, Y. and S. Altman (2002). "Partial reconstitution of human RNase P in HeLa cells between its RNA subunit with an affinity tag and the intact protein components." Nucleic acids research 30(17): 3706-3711.

Lingner, J. and T. R. Cech (1996). "Purification of telomerase from Euplotes aediculatus: requirement of a primer 3' overhang." Proceedings of the National Academy of Sciences of the United States of America 93(20): 10712-10717.

Lowary, P. T. and O. C. Uhlenbeck (1987). "An RNA mutation that increases the affinity of an RNA-protein interaction." Nucleic acids research 15(24): 10483-10493.

Misaki, A., M. Torii, T. Sawai and I. J. Goldstein (1980). "Structure of the Dextran of Leuconostoc-Mesenteroides-B-1355." Carbohydrate Research 84(2): 273-285.

Ni, C. Z., R. Syed, R. Kodandapani, J. Wickersham, D. S. Peabody and K. R. Ely (1995). "Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly." Structure 3(3): 255-263.

Peabody, D. S. (1993). "The RNA binding site of bacteriophage MS2 coat protein." The EMBO journal 12(2): 595-600.

Pillai, R. S., C. G. Artus and W. Filipowicz (2004). "Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis." RNA 10(10): 1518-1525.

Rouault, T. A., M. W. Hentze, D. J. Haile, J. B. Harford and R. D. Klausner (1989). "The iron-responsive element binding protein: a method for the affinity purification of a regulatory RNA-binding protein." Proceedings of the National Academy of Sciences of the United States of America 86(15): 5768-5772.

SenGupta, D. J., B. Zhang, B. Kraemer, P. Pochart, S. Fields and M. Wickens (1996). "A three-hybrid system to detect RNA-protein interactions in vivo." Proceedings of the National Academy of Sciences of the United States of America 93(16): 8496-8501.

Scharpf, M., H. Sticht, K. Schweimer, M. Boehm, S. Hoffmann and P. Rosch (2000). "Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex." European journal of biochemistry / FEBS 267(8): 2397-2408.

Srisawat, C. and D. R. Engelke (2001). "Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins." RNA 7(4): 632-641.

Srisawat, C. and D. R. Engelke (2002). "RNA affinity tags for purification of RNAs and ribonucleoprotein complexes." Methods 26(2): 156-161.

Srisawat, C., I. J. Goldstein and D. R. Engelke (2001). "Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures." Nucleic acids research 29(2): E4.

Srisawat, C., F. Houser-Scott, E. Bertrand, S. Xiao, R. H. Singer and D. R. Engelke (2002). "An active precursor in assembly of yeast nuclear ribonuclease P." RNA 8(10): 1348-1360.

Stripecke, R., C. C. Oliveira, J. E. McCarthy and M. W. Hentze (1994). "Proteins binding to 5' untranslated region sites: a general mechanism for translational regulation of mRNAs in human and yeast cells." Molecular and cellular biology 14(9): 5898-5909.

Tuerk, C. and L. Gold (1990). "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." Science 249(4968): 505-510.

van Gelder, C. W., S. I. Gunderson, E. J. Jansen, W. C. Boelens, M. Polycarpou-Schwarz, I. W. Mattaj and W. J. van Venrooij (1993). "A complex secondary structure in U1A pre-mRNA that binds two molecules of U1A protein is required for regulation of polyadenylation." The EMBO journal 12(13): 5191-5200.

Walker, S. C., F. H. Scott, C. Srisawat and D. R. Engelke (2008). "RNA affinity tags for the rapid purification and investigation of RNAs and RNA-protein complexes." Methods in molecular biology 488: 23-40.

Wolff, T. and A. Bindereif (1992). "Reconstituted mammalian U4/U6 snRNP complements splicing: a mutational analysis." EMBO J 11(1): 345-359.

Xiao, S., F. Scott, C. A. Fierke and D. R. Engelke (2002). "Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes." Annu Rev Biochem 71: 165-189.